Computer Models of Bacterial Cells: From Generalized Coarse-Grained to Genome Specific Modular Models

Evgeni V. Nikolaev[1] (en46@cornell.edu), Jordan C. Atlas[2], Michael L. Shuler[1,2]

[1]Department of Biomedical Engineering, [2]School of Chemical and Biomolecular Engineering, Cornell University, Ithaca NY

Summary

While genomics allows the blueprints of life to be read into static networks of thousands of interacting biomolecules, a better understanding of the functioning of such networks is still required. Here we propose a dynamic modeling framework to integrate genomic detail and cellular physiology within functionally complete 'hybrid' bacterial cell models. An initial step in this approach is the development of a whole-cell coarse-grained model which explicitly links DNA replication, metabolism, and cell geometry with the external environment. A hybrid model can then be constructed from chemically-detailed and genome-specific subsystems, called modules, inserted into the original coarse-grained model. To illustrate the modeling principles, the Cornell coarse-grained *E.coli* model, comprised of 36 ODEs, two algebraic equations, and 31 discrete events, is considered.

Modular Approach:

• Construction of generalized coarse grain single cell models of bacteria has been pioneered by the Shuler group [1-6].

• Our coarse-grained hybrid models explicitly links DNA replication to **metabolism**, **cell cycle**, **cell geometry** and **external environment** [2].

• Our models incorporate non-metabolic function such as **control of chromosomal replication** [4].

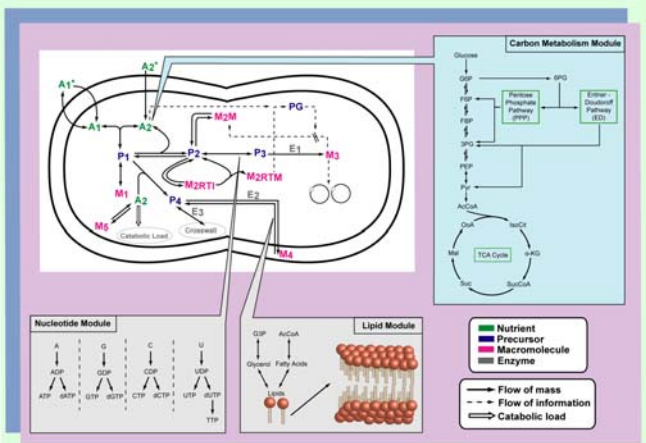


Figure 1. **A schematic representation of the single cell model and the modular approach to cell modeling.** Grey boxes indicate modules that have been implemented (nucleotide metabolism [1] and lipid metabolism [6]). The blue box illustrates a new module under development. Note: Not all reactions and regulation information is depicted. A_1 – ammonium ion, A_2 – glucose, P_1 – amino acids, P_2 – ribonucleotides, P_3 – deoxyribonucleotides, P_4 – membrane precursors, M_1 – protein, M_{2RTI} = immature stable RNA, M_{2RTM} – mature stable RNA, M_3 – DNA, M_4 – cell envelope, M_5 – glycogen, PG – ppGpp, E_1 – enzymes for conversion of P_2 to P_3, E_2 & E_3 – enzymes for cross-wall formation and cell envelope synthesis.

Sample Simulation Results:

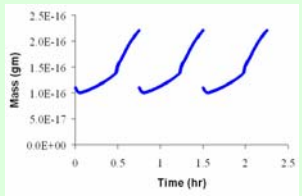


Figure 2. **The mass of ammonium ions (A_1 in Figure 1).**

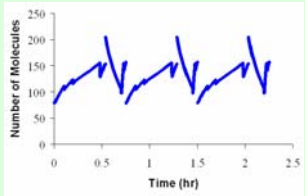


Figure 3. **Free *DnaA-ATP*.**

Stability and Robustness of Hybrid Modular Models:

• Return map P relates the model's state variables Y between consecutive divisions k and k+1, $Y_{k+1} = P(Y_k)$. A stationary cell division cycle corresponds to a fixed point Y_0 of P, $Y_0 = P(Y_0)$, and the cycle's stability can be evaluated from the calculation of the eigenvalues of the linearization $dP(Y_0)$.

• The cycle is stable when *all* eigenvalues are found within the interior of the unit circle in the complex plane.

• These techniques will allow us to generate **bifurcation diagrams** by avoiding *non-*smooth time courses due to discrete cellular events (eg. cell division).

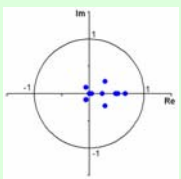


Figure 4. **Eigenvalues of the linearization of return map *P*** The largest *positive* eigenvalue is an important measure of the cell's robustness or biological survivability, while *complex-conjugated* eigenvalues are indicative of potential intracellular oscillatory processes.

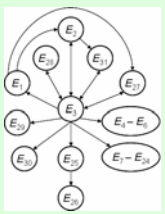


Figure 5. **Events:** completion of DNA methylation (*E*1), transition of replicon state (*E*2), DNA replication initiation (*E*3), changes in dnaA dosage (*E*4) – (*E*6), changes in *rrn*-operon dosage (*E*7) – (*E*24), DNA replication termination (*E*25), cell division (*E*26), the ability of DnaA-ATP (*E*27) and DnaA-ATP (*E*28) to bind high affinity DNA boxes, the ability of DnaA to bind medium affinity boxes (*E*29), the ability of DnaA to bind nonspecific boxes (*E*30), and the ability of DnaA to bind the triggering R5 box within *oriC* (*E*31).

The Minimal Cell Model:

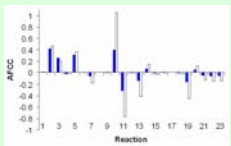
• In related work we are building a **Minimal Cell Model (MCM)**, which uses a "bottom up" approach; the necessary model functions are selected by rationally deciding what machinery a cell needs to live and reproduce [2].

• A minimal cell is a hypothetical cell defined by the **fundamental functions required for life**: It has the minimum number of genes necessary and sufficient for a cell to divide and grow in an optimal environment.

• We have demonstrated that the **modular nature** of this approach by inserting genomically detailed models of nucleotide and lipid metabolism into the coarse grained model [1].

Sensitivity Analysis:

• Using a cell cycle relative phase, amplitudes of the first-order periodic sensitivity functions or averaged flux control coefficients (AFCC) can be ranked to identify important processes and delineate those modules for which additional genomic and chemical detail would be required.



$$AFCC = \frac{\partial \ln(\bar{J})}{\partial \ln(p)}$$

where *AFCC* is the averaged flux control coefficient for the rate parameter p with period - averaged flux $\bar{J}$.

Figure 6. **AFCC values.** Blue and white bars correspond to AFCCs of the specific growth and lipids synthesis rates, respectively. AFCC is defined below the chart. Processes labeled by integer numbers correspond to reactions as follows: 1 – NH_4^+ transport driven by facilitated diffusion, 2 – NH_4^+ transport coupled with membrane energy processes, 3 – glucose transport, 4 – proton leakage, 5 – amino acids synthesis, 6 – amino acids degradation, 7 – free bases synthesis, 8 – free bases degradation, 9 – deoxyribonucleotides synthesis, 10 – cell envelope precursors synthesis, 11 – protein synthesis, 12 – protein degradation, 13 – DNA synthesis, 14 – cross-wall and cell-envelope synthesis, 15 – cell envelope degradation, 16 – glycogen synthesis, 17 – glycogen degradation, 18 –synthesis of enzyme E_1, 19 – M_{2RTI} synthesis, 20 – M_{2RTI} degredation, 21 – M_{2RTM} processing rate, 22 – M_{2M} synthesis, 23 – M_{2M} degradation.

Future Work:

• Develop sensitivity and bifurcation analysis of large scale hybrid differential-algebraic models with discrete events.

•Application of this modeling framework to bifurcation and robustness of cell division cycle and DNA replication initiation, modulated oscillations transmitted between generations ("cell-clock").

•Application of this approach can be used to build genomically specific models for gram-negative bacteria such as *S. oneidensis* (remediation of heavy metal wastes) and *Z. moblis* (fermentations of ethanol from renewable biological energy sources).

References:

[1] M. Castellanos, D.B. Wilson, M.L. Shuler, (2004) *Proceedings of the National Academy of the Sciences.*

[2] S. Browning, M.L. Shuler, (2001) *Biotechnology and Bioengineering.*

[3] M.L. Shuler, (1999) *Journal of Biotechnology.*

[4] S. Browning, M. Castellanos, M.L. Shuler, (2004) *Biotechnology and Bioengineering.*

[5] M. Domach, M.L. Shuler, (1984) *Biotechnology and Bioengineering.*

[6] M. Castellanos, M.L. Shuler, (2006) (*to be submitted)*

[7] E.N. Nikolaev, J.C.Atlas, M.L. Shuler (2006) *Journal of Physics Conference Proceedings* (*submitted*).

Acknowledgements:

Jordan Atlas gratefully acknowledges support from the Department of Energy Computational Science Graduate Fellowship Program of the Office of Science and National Nuclear Security Administration in the Department of Energy under contract DE-FG02-97ER25308.

This work is supported by DOE grant DE-FG02-04ER63806.

Michael L. Shuler acknowledges the support of NYSTAR, the New York State Office of Science, Technology, and Academic Research.